4. Real-Time Correction of Spatially Nonuniform Bias in Radar Rainfall Data Using Rain Gauge Measurements

ABSTRACT

A procedure for real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements is described. Developed to complement the existing gauge-based bias correction procedures used in the National Weather Service (NWS), the proposed procedure is a generalized local bias estimator that may be used under varying conditions of rain gauge network density and types of rainfall. To arrive at the procedure, the correction problem is formulated as space-time estimation of radar and bin-averaged gauge rainfall from radar rainfall data and rain gauge measurements, respectively, at all hours up to and including the current hour. The estimation problem is then solved sub-optimally via a variant of exponential smoothing. To evaluate the procedure, parameter estimation and true validation were performed using hourly radar rainfall and rain gauge data from the Arkansas-Red Basin River Forecast Center (ABRFC) area. The results indicate that the proposed procedure is generally superior to mean field bias correction, and that the improvement is particularly significant in the cool season.

4.1 Introduction

Since the introduction of weather radars in operational hydrologic forecasting, real-time correction of systematic biases in radar rainfall data using rain gauge measurements has been widely recognized as one of the most important steps in radar-based rainfall estimation (Ahnert et al. 1986, Collier 1986, Smith and Krajewski 1991). Since the first deployment of the Weather Surveillance Radar - 1988 Doppler version (WSR-88D) in 1991, two approaches have emerged for gauge-based bias correction in the National Weather Service (NWS); mean-field bias correction (Ahnert et al. 1986, Smith and Krajewski 1991, Anagnostou et al. 1998, Seo et al. 1999) and local bias correction (Bill Lawrence, personal communications). The former, in one formulation or another (Smith and Krajewski 1991, Seo et al. 1997, Seo et al. 1999), has been in operation since 1993 at a number of River Forecast Centers (RFC) as a part of the software application known as Stage III (Hudlow 1988, Fulton 1998). The latter (see Appendix A) has been in operation at the Arkansas-Red Basin RFC (ABRFC) since 1996 as a part of the software application known as Process 1 (P1).

Due to significant differences in data processing steps before and after bias correction between Stage III and P1, it is difficult to rigorously compare the two bias correction procedures based on evaluation of the operationally produced Stage III and P1 products (Young et al. 2000). Nevertheless, operational experience does confirm the long-held expectations that, though effective in removing biases that are spatially uniform over the entire radar umbrella due, e.g., to lack of radar calibration and inappropriate Z-R parameters, mean field bias correction is of limited utility (and, in some situations, counterproductive) if biases are spatially nonuniform (Seo et al. 1999, see also Section 8), and that local bias correction is very effective in gauge-rich areas, particularly in widespread stratiform rainfall in the cool season (Bill Lawrence, personal

communications).

Despite the generally positive operational experience at ABRFC, implementation of the local bias correction procedure used in P1 in other parts of the country, particularly those that are gauge-poor, could not be pursued rigorously because of a number of limitations inherent in the procedure. For example, the spatial interpolation procedure does not take into account either the sampling differences between the two sensors or the uncertainties in the bias estimate due to lack of sample size. Consequently, the bias estimates are susceptible to large sampling errors. Also, the spatial interpolation is based solely on the triangular geometry of the rain gauge network (see Appendix A), and hence does not account for either the spatial scale of rainfall or the attendant dependence of sampling errors on the spatial configuration of the gauge network. Consequently, the procedure is susceptible to unevenness in and sparsity of the rain gauge network.

The purpose of this paper is to describe a new procedure for real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. Based on the operational experience with the existing procedures summarized above, the proposed procedure is intended to be a generalized local bias estimator that may be used under varying conditions of gauge network density and types of rainfall. Pending further evaluation, the procedure is to be implemented operationally at the River Forecast Centers (RFC) across the country as a component in the RFCWide Multisensor Precipitation Estimator (RFCWide MPE) (Breidenbach et al. 2000) of the Advanced Weather Interactive Processing System (AWIPS).

It is noted here that the formulation of the proposed procedure, as will be seen, may not necessarily be the most appropriate for dealing with range-dependent biases due to nonuniform vertical profile of reflectivity (VPR) such as that due to bright band enhancement. Because they originate from the reflectivity morphology of precipitating clouds as observed through the sampling geometry of radar, range-dependent biases should be corrected first by the use of volume scan reflectivity data (Andrieu and Creutin 1995, Seo et al. 2000) before correction of residual biases may be attempted using rain gauge measurements. We note here that work is under way to implement a procedure for real-time correction of range-dependent bias due to nonuniform VPR (Seo et al. 2000) in the WSR-88D Open Radar Product Generator (ORPG) (Saffle and Johnson 2000) in the very near future. In the interim, the proposed procedure is to serve as an all-purpose algorithm for correcting, to the extent possible, any biases that are spatially nonuniform regardless of range dependency (i.e. except for those due to beam overshooting: see Section 2).

The organization of this paper is as follows. Section 2 describes the problem of local bias correction. Section 3 formulates the correction problem as space-time estimation. Section 4 describes the procedure used to solve the estimation problem. Section 5 describes the radar and rain gauge data used for parameter estimation and validation. Section 6 describes how the parameters are estimated via sensitivity analysis. Section 7 describes the validation experiment. Section 8 summarizes the results. Section 9 provides conclusions and future research recommendations.

4.2 Problem Description

We define the local bias at the radar bin centered at location u_0 at hour k, β_{0k} , as follows:

$$\beta_{0k} \equiv G_{0k} / R_{0k} \tag{1}$$

where G_{0k} and R_{0k} denote the hourly accumulations of bin-averaged gauge and radar rainfall (mm), respectively, at the k-th hour at the radar bin centered at u_0 . The gauge rainfall, G_{0k} , in Eq.(1) may be expressed as:

$$G_{0k} = \|A_0\|^{-1} \int_{A_0} P_k(u) du$$
(2)

where $\|A_0\|$ denotes the area of the radar bin centered at u_0 , and $P_k(u)$ denotes the hourly gauge rainfall (mm) at location u in A_0 (i.e. if a gauge existed at that location) at the k-th hour. For the WSR-88D Digital Precipitation Array (DPA), also known as the Hourly Digital Precipitation (HDP), product (Klazura and Imy 1993), the size of the radar bin is approximately $4x4 \text{ km}^2$ in mid-latitudes. Note that, as defined in Eq.(2), G_{0k} may be considered measurable only if there exists a sufficiently large number of gauges within A_0 to capture the micro-scale variability of rainfall (Ciach and Krajewski 1999).

The radar rainfall, R_{0k} , in Eq.(1) may be expressed as:

$$R_{0k} = \int_{T} Q_0(t) dt$$
 (3)

where T denotes the duration of an hour and $Q_0(t)$ denotes the bin-averaged radar rain rate (mm hr⁻¹) at time t at the bin centered at location u_0 . In Eq.(3), it is assumed that $Q_0(t)$ is free of spatial sampling errors. In reality, however, radar rainfall data are subject to various sources of such errors: 1) at far ranges, (bin-averaged) radar rainfall estimates may be based only on a few or even less (polar) data points, 2) the navigation of the beam may be deviating from that in the standard atmosphere due to strong temperature and/or moisture gradient (Doviak and Zrnic 1984), and 3) rainfall aloft as observed by the radar may be advecting significant distances before reaching the ground (Wilson and Brandes 1979). Even if there are no spatial sampling errors, the radar rainfall estimate of R_{0k} in Eq.(3) is always subject to temporal sampling errors because, with a scanning radar, the temporal integration in Eq.(3) can only be approximated by a discrete summation (Fabry et al. 1994).

Because the gauge measurement of G_{0k} is subject to spatial sampling errors due to microand macro-scale variabilities of rainfall, and the radar estimate of R_{0k} is subject to spatial and temporal sampling errors from various sources as described above, β_{0k} in Eq.(1) is not a deterministic quantity and, as such, can only be estimated in a statistical sense. Given that the purpose of bias correction is to render the adjusted radar rainfall estimate unbiased against the bin-averaged gauge rainfall at least in the mean sense, the simplest form that a local bias estimator may assume is the following:

$$\beta_{0k}^* = \mathbb{E}[G_{0k}] / \mathbb{E}[R_{0k}] \tag{4}$$

where the asterisk signifies that the variable superscripted is only an estimate and E[] denotes the

expectation operator. The bias-corrected radar rainfall estimate at hour k at the bin centered at u_0 is then given by the product of ${\beta_{0k}}^*$ and r_{0k} , where r_{0k} denotes the raw radar rainfall estimate at hour k at the bin centered at u_0 .

As is, however, Eq.(4) is not wholly appropriate for bias correction of radar rainfall data because it reflects not only the bias in radar rainfall estimates given that the radar successfully detected rainfall (the first term in Eq.(5) below) but also that in radar detection of rainfall (the second term in Eq.(5)):

$$\beta_{0k}^{*} = \frac{E[G_{0k} \mid G_{0k} > 0] \ Pr[G_{0k} > 0]}{E[R_{0k} \mid R_{0k} > 0] \ Pr[R_{0k} > 0]}$$
(5)

where Pr[] denotes the probability of occurrence of the event bracketed. Consequently, when and where beam overshooting does occur, bias correction based on Eq.(4) will necessarily overcorrect the radar rainfall estimates that are free of beam overshooting to make up for the amount of rainfall unaccounted for by radar due to beam overshooting. Because correcting biases in radar detection of rainfall is an extremely difficult problem for which no practical solutions currently (if ever) exist, the only operationally viable option is to avoid dealing with such biases altogether by limiting the use of radar rainfall data to within the 'effective coverage' of radar where $Pr[G_{0k}>0] \approx Pr[R_{0k}>0]$ holds (i.e. where radar can consistently 'see' rainfall: see Section 5). Under this restriction, we may drop the second term in Eq.(5) to write:

$$\beta_{0k}^{*} = E[G_{0k} \mid G_{0k} > 0] / E[R_{0k} \mid R_{0k} > 0]$$
(6)

Because the bias estimate, β_{0k}^* , is expressed as a ratio, particular care must be taken to avoid inadvertently introducing biases in estimation of the conditional expectations in Eq.(6). In the next section, we describe how we arrive, in light of the above observations, at the particular form and choice of the estimator used in the proposed procedure.

4.3 Estimation Approach

Because bias estimation as defined in Eq.(6) is a problem of inferring first-order moments of two independent observations of rainfall, it behooves to estimate them independently of each other based exclusively on gauge and radar rainfall data, respectively. In this way, estimation of mean gauge rainfall does not require any knowledge, a priori or otherwise, about the mean radar rainfall, and vice versa. Accordingly, to estimate $E[G_{0k} | G_{0k} > 0]$ for an ungauged bin, some sort of spatial averaging of neighboring rain gauge measurements is necessary. Furthermore, because the number of rain gauge measurements available in real time is very often very small, some sort of temporal averaging may also be necessary to improve the statistical reliability of the bias estimate (even if it may mean having to compromise the unbiasedness to some degree: see Section 8). With these observations, we expand Eq.(6) to write:

$$\beta_{0k}^{*} = \frac{E[G_{0k} \mid G_{0k} > 0, G_{ij}, i=1,...n_{j}, j=1,...,k, G_{ij} > 0]}{E[R_{0k} \mid R_{0k} > 0, R_{ij}, i=1,...n_{j}, j=1,...,k, R_{ij} > 0]}$$
(7)

where G_{ij} and R_{ij} , $i=1,...n_j$, j=1,...,k, denote the gauge and radar rainfall, respectively, in the neighborhood of u_0 at the current and the preceding hours, and n_j denotes the number of positive radar-gauge pairs available at the j-th hour. Adopting filtering notations for brevity, we denote the denumerator and denominator in Eq.(7) as $g_{0k|k}$ and $r_{0k|k}$, respectively.

In Eq.(7), we could have dropped $\{G_{ij}>0\}$ and $\{R_{ij}>0\}$ from the conditioning sets: i.e., rather than using only positive rainfall data, use both zero and positive rainfall data to estimate the conditional mean. Such a conditioning, however, requires that both the inner variability and the intermittency of rainfall be taken into account, which not only complicates the estimation problem significantly but, more importantly, may also introduce biases in β_{0k}^* due to additional statistical parameterization necessary to handle intermittency (Seo 1998). Note also in Eq.(7) that indexing of R_{ij} is identical to that of G_{ij} : i.e., estimate the conditional mean of radar rainfall using only those data collocated with rain gauge measurements even if there may be many other radar rainfall data available in the immediate vicinity of the bin centered at u_0 . Such selective sampling of radar rainfall data is necessary to avoid introducing sampling biases in β_{0k}^* due to spatial variability of rainfall.

To estimate the conditional expectations in Eq.(7), it is tempting to consider optimal estimation techniques that combine a priori mean from climatology and sample mean from current observations (see, e.g., Smith and Krajewski 1991). In that way, the bias estimate may be rendered, in some objective fashion, to be more reflective of the climatology in gauge-sparse areas and of the current observations in gauge-dense areas. For such a purpose, at least two techniques are available; space-time kriging (Rouhani and Myers 1990, Rouhani and Wackernagel 1990) based on simple cokriging (Joumel and Huijbregts 1978) and multivariate Kalman filtering (see, e.g., Bras and Rodriguez-Iturbe 1985). (The two are equivalent under certain conditions: see Seo and Smith 1996.) Unfortunately, for purposes of real-time bias estimation (as opposed to climatological or post analysis), neither technique could be considered viable. The reason is that, in the presence of large uncertainties in key micro-physical parameters such as the Z-R and in the calibration of the radar, it is extremely difficult (if not practically impossible) to come up with unbiased and informative a priori estimates of conditional mean of gauge and radar rainfall. As such, we considered in this work only those techniques that do not assume any a priori knowledge about the mean.

4.4 Estimation Procedure

Given the above observations, we sought as simple and parsimonious as possible a space-time estimator that does not assume any a priori knowledge about the mean. The technique chosen here is a variant of exponential smoothing (Schweppe 1973) which, in its original form, has been used in the mean field bias correction procedure (Seo et al. 1999). The particular formulation used in this work solves the following constrained weighted least-squares minimization for R_{0k} (that for G_{0k} is completely analogous: see below):

$$\min_{j=1}^{k} J_{r} = \sum_{j=1}^{k} e^{-(k-j)/\alpha} \left[Z_{rj} - H_{j} R_{0k} \right]^{T} \Psi_{rj}^{-1} \left[Z_{rj} - H_{j} R_{0k} \right]$$
(8)

subject to
$$R_{0k} = R_{0(k-1)}$$
 (9)

where k denotes the current hour, α denotes the memory span (hr), Z_{rj} denotes the $(n_j)x(1)$ measurement vector of radar rainfall data at hour j, $[r_{1j},r_{2j},...,r_{(nj)j}]^T$, H_j denotes the $(n_j)x(1)$ unit vector, $[1,1,...,1]^T$, and Ψ_{rj} denotes the $(n_j)x(n_j)$ error covariance matrix at hour j. In the above, n_j denotes the number of data points at hour j. Throughout this paper, 'r' and 'g' signify that the variable subscripted is associated with radar and gauge rainfall, respectively.

To obtain the recursive solution for R_{0k} , we make use of the fact that, without the exponential decay term in Eq.(8), the recursive solution to the above minimization problem is identical to the Kalman filter solution (Schweppe 1973) under the system model of Eq.(9) and the observation model of:

$$Z_{rk} = H_k R_{0k} + V_{rk}$$
 (10)

where V_{rk} denotes the $(n_k)x(1)$ error vector with $E[V_{rk}]=0$ and $E[V_{rk}V_{rk}^{\ T}]=\Psi_{rk}$. Then, by rewriting the Kalman filter solution first into a non-recursive form while replacing $\Psi_{rk}^{\ -1}$ with $e^{-(k-1)/\alpha}$ $\Psi_{rk}^{\ -1}$, l=1,...,k-1, and then again into a recursive form (see Schweppe 1973 for details), we obtain the following exponential smoothing solution for R_{0k} , $r_{0k|k}$, and a measure of uncertainty associated with it, $\Sigma_{r0k|k}$:

$$r_{0k|k} = r_{0(k-1)|(k-1)} + \sum_{r_{0k|k}} H_k^T \Psi_{rk}^{-1} (Z_{rk} - H_k r_{0(k-1)|(k-1)})$$
(11)

$$\Sigma_{r0k|k}^{-1} = e^{-1/\alpha} \sum_{r0(k-1)|(k-1)}^{-1} + H_k^T \Psi_{rk}^{-1} H_k$$
(12)

where the initial conditions are given by $r_{00|0}=0$ and $\Sigma_{r00|0}^{-1}=0$ (i.e. completely uninformative prior). Note that, because errors are not explicitly modeled in exponential smoothing in favor of parsimony, $\Sigma_{r0k|k}$ is only a relative measure of uncertainty and may not in general be interpreted as the error variance associated with $r_{0k|k}$.

The trick that renders Eqs.(11) and (12) a space-time estimator is to define the error vector, V_{rk} , in Eq.(10) not as measurement errors as invariably modeled in Kalman filtering but as first-order differences of rainfall in space, $[R_{1k}-R_{0k},R_{2k}-R_{0k},...,R_{(nk)k}-R_{0k}]^T$. In this way, the spatial variability of rainfall may be accounted for in the temporal smoothing process of Eqs.(11) and (12) while the actual measurement errors in the rainfall data are, for the sake of parsimony, ignored. Then, the ij-th entry in the error covariance matrix, Ψ_{rk} , is given by:

 ψ_{rijk}

$$= Cov[R_{ik} - R_{0k}, R_{ik} - R_{0k}]$$
(13a)

$$= Var[R_k] + Cov[R_{ik}, R_{ik}] - Cov[R_{ik}, R_{0k}] - Cov[R_{ik}, R_{0k}]$$
(13b)

$$= \gamma [R_{ik}, R_{0k}] + \gamma [R_{ik}, R_{0k}] - \gamma [R_{ik}, R_{ik}]$$
(13c)

where Cov[] and γ [] denote the covariance and the semi-variogram (Journel and Huijbregts 1978), respectively. It can be shown then (see Appendices B and C) that the resulting space-time estimation amounts to temporally weight-averaging ordinary kriging estimates from the current and previous hours according to the 'age' of the estimates, and that the resulting estimate, $r_{0k|k}$, is unbiased in the mean sense. The exponential smoothing procedure used to estimate the binaveraged gauge rainfall at hour k centered at u_0 , G_{0k} (see Eq.(2)), is completely analogous to that for R_{0k} . The only difference is that, because we are estimating bin-averaged gauge rainfall using measurements of point gauge rainfall, the exponential smoothing procedure amounts to temporal averaging of ordinary block kriging estimates rather than that of ordinary punctual kriging estimates (see Appendix D).

The exponential smoothing procedure described above yields the unbiased minimum-errorvariance solution if the (space- and time-varying) memory span, α , is optimal. By comparing the exponential smoothing solution to the Kalman filter solution, it can be easily shown that specifying α amounts to specifying the ratio of the model uncertainty to the filtered variance in the Kalman filter solution. Because we are using exponential smoothing as a simple and parsimonious (albeit sub-optimal) alternative to Kalman filtering, we are not necessarily interested here in estimating the optimal α . (It would entail performing rigorous, and probably too difficult to be practical, estimation of space- and time-varying parameters, comparable to that for distributed-parameter Kalman filtering (see, e.g., Tzafestas 1978)). Instead, we estimate in this work a sub-optimal α based on the parallel estimation strategy employed in the mean field bias correction procedure (Seo et al. 1999). In short, the strategy amounts to performing exponential smoothing at multiple values of memory span, α (corresponding, e.g., to hourly, daily, weekly, monthly, seasonal, yearly, etc., scales of temporal aggregation), and selecting, among the solutions for which the relative measure of uncertainty associated with β_{0k}^* , $\Sigma_{\beta 0k|k}$, is smaller than some prescribed threshold, Σ_{β} , that with the largest $\Sigma_{\beta 0k|k}$ as the 'best' bias estimate. In the above, $\Sigma_{80k|k}$ may be estimated under the assumption that R_{0k} , G_{0k} , and β_{0k} are lognormally distributed (see Seo et al. 1999 for details). The rationale behind the strategy is extremely simple: if the bias estimate based only on the data for the current hour is already of acceptable statistical reliability, it is your best estimate; if not, successively include the most recent data in the estimation until the resulting estimate is of acceptable statistical reliability.

Because it is applied to a relative measure of uncertainty, the threshold $\Sigma_{\beta t}$ is a rather difficult parameter to work with, particularly in an operational setting. For this reason, rather than applying a threshold on $\Sigma_{\beta 0k|k}$, here we apply one on the number of effective positive radargauge pairs within the estimation neighborhood of u_0 , referred to herein as the radius of influence (see Section 6). We denote this threshold as N_t . The effective number of positive radar-gauge pairs is updated recursively in a manner completely analogous to Eq.(12) (see also Seo et al. 1999):

$$N_{0k|k} = e^{-1/\alpha} N_{0(k-1)|(k-1)} + N_k$$
 (14)

4.5 Data

To estimate the parameters and to evaluate the procedure, true validation is performed using hourly radar and rain gauge data from the ABRFC area. In this section, we describe the data sets used and how they are prepared for the experiments.

a. Radar Data

The radar data used are the WSR-88D Digital Precipitation Array (DPA) products (Klazura and Imy 1993) from the WSR-88D sites within and in the vicinity of the ABRFC area (see Fig 1). The period covered is May 1, 1996, through April 30, 1999, encompassing three warm (May through September) and cool (October through April) seasons. The first year (May of 1996 through April of 1997) and the last two years (May of 1997 through April of 1999) served as the parameter estimation and the validation periods, respectively. Because local bias correction is to be performed over the entire RFC area in a spatially continuous manner (as opposed to one radar umbrella at a time as in mean field bias correction), it is necessary to mosaic, for each hour, all available DPAs from all WSR-88D sites over the entire RFC area of 335x159 HRAP². An HRAP bin in the area is approximately 4 km a side. The steps involved, which are identical to those used in RFCWide MPE (Breidenbach et al. 2000), are briefly described below.

- 1) For each radar site, delineate the 'effective coverage' of radar for cool and warm seasons based on long-term statistics of DPA products, such as probability of detection of rainfall and mean rainfall. The details are beyond the scope of this paper, and the further interested reader is referred to Breidenbach et al. (1999). The effective coverage thus obtained represents the area within the nominal range of 230 km where the radar can consistently 'see' precipitation so that $Pr[G_{0k}>0] \approx Pr[R_{0k}>0]$ holds (see Eq.(5)).
- 2) For each radar site, calculate, for each azimuth-range bin, the height (km above MSL) of the axis of the beam used to obtain the rainfall estimate in the DPA product. Re-map the resulting height field onto the HRAP grid by averaging all polar data points within each HRAP bin.
- 3) For each HRAP bin in the RFC area, check if the bin lies within the union of the effective coverages of all radars for which DPA products were available for that hour. If it does, choose the radar rainfall estimate sampled at the lowest height as the 'best' for that bin for that hour. If not, no radar rainfall estimate is available at that bin for that hour.

The motivation for the above mosaicking strategy is to seek radar rainfall estimates from the 'lowest unobstructed beam' (O'Bannon 1997) in areas of coverage overlap. For further details,

the reader is referred to Breidenbach et al. (1999, 2000).

b. Rain Gauge Data

The rain gauge data used are the hourly data that were operationally available to the RFC during the period (see Fig 1 for locations). Real-time rain gauge data are subject to numerous sources of error; mechanical malfunction, electronic malfunction (including transmission errors), (lack of) exposure, wind, freezing and thawing, biological contamination, inaccurate georegistration, etc. (see also Steiner et al. 1999). As such, in addition to automatic checks, the rain gauge data should ideally undergo visual spatio-temporal checks with the aid of radar rainfall data (and, preferably, satellite data as well). Visual examination of large quantities of data, however, is an extremely labor-intensive task and was well beyond the scope of this work. As a compromise, we performed only the following very simple, but stringent, statistical check in this work:

- 1) for each gauge location, extract the time series of hourly gauge and radar rainfall for the cool and warm seasons,
- 2) for each season and for each gauge location, calculate the probabilities of detection of rainfall, the indicator and conditional cross-correlation coefficients between gauge and radar rainfall, and the ratio of the sum of gauge rainfall to that of radar rainfall, referred to herein as the gauge-specific bias,
- 3) for each season, apply thresholds to the statistics calculated above to screen out 'apparently bad' gauges, and
- 4) for each season and for each 'apparently good' gauge, screen out apparent outliers in the gauge measurements by applying a threshold to the absolute difference between the gauge and the gauge-specific bias-adjusted radar rainfall.

In Steps 1 through 3, whenever the integrity of the gauge was in doubt, we chose to err on the side of caution and threw out the entire data set from that gauge for that season. Hence, it is possible that some of the natural variabilities of rainfall may have been lost and/or the sampling errors in the rain gauge measurements (see Section 2) may be under-represented in the surviving data set.

4.6 Parameter Estimation

The proposed procedure, as formulated in Section 4, has as many as ten adaptable parameters (excluding the number of memory spans and their values, which are considered to be fixed parameters). Eight of them pertain to the semi-variogram models of gauge and radar rainfall; the types of the model (limited to spherical, gaussian, and exponential in this work), the sills, C_g and C_r (mm²), the nugget effects, μ_g and μ_r (mm²), and the correlation scales, L_g and L_r (HRAP). The other two are the radius of influence, RI (HRAP), and the threshold for the effective number of positive radar-gauge pairs, N_r (see Eq.(14)).

Ideally, the semi-variograms of gauge and radar rainfall should be estimated locally at every hour, so that the spatio-temporal variations in the spatial variability of rainfall may be accounted for in the estimation process. In reality, however, such fine-scale modeling of semi-

variograms is rarely (if ever) possible because of lack of rain gauge measurements. As such, we use the climatological semi-variograms instead, i.e., the semi-variograms estimated from all data points within the entire area of interest and within the entire period of record available (Bastin et al. 1984, Lebel et al. 1987, Lebel and Laborde 1988, Seo 1996). Fig 2 shows the climatological semi-variograms of radar and gauge rainfall for the cool and warm seasons over the ABRFC area (see Fig 1) as estimated from the one-year data in the parameter estimation period. In the figure, the markers '1' through '8' denote the directional experimental semi-variograms along the angles of 0, 26.6, 45, 63.4, 90, 116.6, 135, and 153.4 degrees (counterclockwise from due east). Also shown in each figure is the best weighted least-squares fit, whose parameters are summarized in Table 1. It is important to note that, in estimating the experimental semi-variograms, we used only the collocated radar-gauge pairs. The reason for this selective sampling of radar data is that the performance of the bias correction procedure is rather sensitive to the relative magnitude of L_g and L_r (see Section 8). As such, sampling biases in the experimental semi-variograms of gauge and radar rainfall must be avoided, even if it may mean reducing the reliability of the experimental semi-variogram of radar rainfall due to reduced sample size.

Use of the climatological semi-variograms implies that the variances of radar and gauge rainfall are modeled as space- and time-invariant (i.e. intra-seasonally). Then, it can be easily shown that $r_{0k|k}$ (see Eq.(11)) and $g_{0k|k}$ are independent of C_g and C_r , and therefore we may, without any loss of generality, replace the nugget effects, μ_g and μ_r , with the normalized nugget effects, $\mu_g/C_g\equiv(n_g)$ and μ_r/C_r ($\equiv n_r$), respectively, set $C_g\equiv C_r\equiv 1$, and define Ψ_{rk} and Ψ_{gk} in Eqs.(13) and (D2), respectively, as error correlation (as opposed to covariance) matrices. Setting sills to a constant, however, leaves quantification of the relative magnitude of the sampling uncertainty in the Fisher estimate (see Eq.(B2)) completely at the mercy the Fisher error variance (see (B1)), even though the latter may not be a realistic representation of the effect of sample size on the sampling uncertainty. As such, we express C_g and C_r as an explicit function of the sample size in this work to explicitly account for the uncertainty due to lack of sample size:

$$C_g = C_r = 1/(N_k + 1)^v$$
 (15)

where N_k denotes the number of collocating positive radar-gauge pairs within the radius of influence, RI (HRAP), at hour k, and v is a nonnegative number. The general form of the dependence of sampling uncertainty on sample size as expressed in Eq.(15) is supported by Rodriguez-Iturbe and Mejia (1974) and used by Smith and Krajewski (1991). The fixed parameter v is estimated via sensitivity analysis similar to those performed for estimation of free adaptable parameters (see Section 8). The analysis indicates that, over [0,2], a larger (smaller) v yields a somewhat smaller mean error (root mean square error) in the bias-corrected radar rainfall estimates but only at the expense of a somewhat larger root mean square error (mean error), and that v=1 is a reasonable compromise. As such, v was fixed at unity throughout the rest of this work. It is worth noting that, with v=1, the local bias correction procedure is reduced to the mean field bias correction procedure (Seo et al. 1999) if both the radar and gauge rainfall are white-noise in space.

The use of climatological semi-variograms also implies that the second-order spatial statistics of rainfall are assumed to be space- and time-invariant. In reality, however, the

statistics vary in space and time. Also, the majority of the data points used in their estimation comes from light to medium rainfall events. As such, the climatological semi-variograms are not in general reflective of the spatial structure of rainfall in heavy rainfall events. For these reasons, the local bias-corrected estimates based on the climatological semi-variograms are in general subject to rainfall amount-dependent (i.e. conditional) biases, even though they may be free of mean bias (i.e. bias averaged over all ranges of rainfall amount). As such, some form of sensitivity analysis or parameter optimization (using the climatological semi-variogram parameters as the initial guess) is generally necessary, not only to pin down the optimal parameter settings but also to ascertain the nature of the conditional bias and its dependence on parameter settings. In Section 8, we describe how a series of sensitivity analysis is performed for such purposes.

The radius of influence, RI (HRAP), represents the length scale at which the spatial process may be considered locally homogeneous. It defines the circular (under the isotropy assumption) neighborhood of estimation centered at the bin where the bias estimate is sought. All gauge measurements and collocating radar rainfall data within the circle are then considered as neighbors of G_{0k} and R_{0k} (see Eq.(7)), and are used in the estimation of $g_{0k|k}$ and $r_{0k|k}$, respectively. Determination of this scale based on local homogeneity considerations, however, is neither straightforward (Journel and Huijbregts 1978) nor necessarily desirable as explained below. In gauge-sparse areas, it is preferable to have as large a radius of influence as possible, regardless of the scale at which local homogeneity may or may not actually hold, so that the bias estimates do become available over as large an area as possible. Hence, as long as the performance of the bias correction procedure does not deteriorate and the computationally burden remains bearable, RI should be set to a largest possible value. As such, RI is treated as a free parameter in this work and estimated via sensitivity analysis (see Section 8).

4.7 Validation

To evaluate the proposed procedure, true validation is performed using the data from the two-year validation period. The steps involved are as follows:

- 1) specify the density of the gauge network,
- 2) divide the operational gauge network (see Fig 1) into two; the estimation network, which has the density specified in Step 1, and the validation network, which is given by the operational network minus the estimation network.
- 3) for each hour of the two-year validation period, perform mean field and local bias correction at rain gauge locations in the validation network using the gauge measurements from the estimation network and the collocated radar rainfall data,
- 4) collect the raw, mean field bias-corrected, and local bias-corrected radar rainfall estimates at the gauge locations in the validation network and the (now revealed) rain gauge measurements, and
- 5) calculate performance measures.

In Step 1, two densities are assumed for the estimation network; one-half and one-fourth of that of the full operational gauge network. They are referred to as the $\frac{1}{2}$ and $\frac{1}{4}$ networks, and have 230 and 115 gauges, respectively, that are randomly selected from the operational network

(see Fig 1). In Step 4, the mean field bias correction procedure used is that of Seo et al. (1999), in which the number of memory spans used is 10, and their settings are 1, 2, 4, 8, 16, 32, 64, 128, 256, and 10⁶ (hr). The same number and settings are also used for local bias correction. The threshold for the effective number of radar-gauge pairs used in the mean field bias correction procedure (i.e., the equivalent of N_t in this work) was 16 for both seasons based on Seo et al. (1999). The performance measures used in Step 5 are the same as those used in Seo et al. (1999); the ratio of the sum of gauge rainfall to that of radar rainfall (RATIO), the root mean square error of radar rainfall (RMSE) (or, alternatively, the mean square error (MSE)), the cross-correlation coefficient between gauge and radar rainfall (CORR), and the maximum errors of under- and overestimation by radar rainfall (MAXEU and MAXEO, respectively).

Because local bias-corrected radar rainfall is an areal average whereas rain gauge data are point measurements, the rain gauge measurements from the validation network should ideally be converted to areal averages before calculating the performance measures. Although such a conversion via, e.g., ordinary block kriging, should reduce sampling errors, it is bound to increase estimation errors. As such, no attempts were made in this work to convert the rain gauge measurements to bin-averaged gauge rainfall estimates. Hence, all results presented in the next section are based on validation of areally-averaged (radar) rainfall against point (gauge) rainfall. The consequence of ignoring this difference in spatial sampling scale of rainfall depends on the performance measure. Because it is based on long-term accumulations, RATIO should not be affected at all. For MSE, it may be shown (see Appendix E) that, even though the absolute magnitude of MSE is subject to sampling errors due to micro-scale variability of rainfall, that of reduction in MSE is not, and that percent reduction in MSE based on point validation is a conservative estimate of that based on areal validation. (The above results do not, however, hold for RMSE.) Because point rainfall has larger variability than the areally averaged, MAXEO and MAXEU based on point validation should also be conservative estimates of those based on areal validation.

4.8 Results

Because the number of parameters involved is rather large, it is difficult to assess the sensitivity of the proposed procedure to all parameters through one large analysis. Instead, here we performed a series of smaller sensitivity analyses, each involving fewer parameters. The first set of parameters examined was RI and N_t . Using the climatological semi-variogram parameters for the types of the variogram model for radar and gauge rainfall, n_g , n_r , L_g , and L_r , true validation was performed at various combinations of parameter settings of RI and N_t . The contour plots of RATIO, RMSE, CORRE, MAXEO, and MAXEU of the local bias-corrected estimates were then examined as functions of RI and N_t . They indicate that RI has a rather straightforward influence on the correction procedure: a larger RI improves the performance, but, beyond RI \approx 60 (HRAP), the improvement is negligible. As such, we fixed RI at 60 (HRAP) (corresponding to approximately 240 km) throughout the rest of this work. Unless mentioned otherwise, all results presented in this section are based also on the $\frac{1}{2}$ gauge network (see Section 7).

Fig 3 shows the five performance measures for the hourly local bias-corrected estimates based on the climatological semi-variogram parameters. For comparison, those for the raw radar

rainfall and mean field bias-corrected estimates are also shown. Figs 3a and 3b show RATIO and RMSE of the raw (denoted as 'r'), mean field bias-corrected (denoted as 'm'), the local biascorrected (denoted as '1' through '6') radar rainfall estimates for cool and warm seasons, respectively. The markers '1' through '6' signify that the local bias-corrected estimates are based on N, settings of 2, 4, 8, 16, 32, and 64, respectively (see Section 4). The perfect estimates would locate (RATIO,RMSE) at (1,0) in the figures, where for reference the line of perfect RATIO (i.e. of unity) is also shown. Fig 3c shows CORRE. The perfect estimates would locate (cool-season CORRE, warm-season CORRE) at (1,1). Figs 3d and 3e show MAXEU (defined as the negative of min{estimate-gauge rainfall}) and MAXEO (defined as max{estimate-gauge rainfall}) for cool and warm seasons, respectively. The perfect estimates would locate (MAEXEU, MAXEO) at (0,0). Fig 3 may be summarized as follows: 1) local bias correction based on climatological semi-variograms is very effective in removing the mean bias (i.e. bias averaged over all amounts of rainfall), 2) temporal smoothing is important in the warm season, particularly to reduce RMSE and MAXEO, 3) N_t settings of 8 and 32 provide appropriate levels of temporal smoothing for estimation of hourly rainfall in cool and warm seasons, respectively, and 4) at such settings of N₁, the mean performance (i.e. performance averaged over all ranges of rainfall amount) of the proposed procedure based on the climatological semi-variograms is generally superior, particularly in the cool season, to that of the mean field bias correction procedure (Seo et al. 1999).

Fig 3, however, pertains mostly to the performance of the procedure averaged over all ranges of rainfall depth. As noted in Section 6, local bias-corrected estimates based on climatological semi-variograms are subject to conditional (i.e. on the amount of rainfall being estimated) biases. Indeed, examination of the scatter plots of daily rainfall amounts between the local bias-corrected estimates and the gauge measurements indicates that the local bias-corrected estimates associated with Fig 3 tend to significantly underestimate large rainfall amounts (and, conversely, overestimate small rainfall amounts). To examine the sensitivity of the conditional (i.e. on the amount of rainfall being estimated) performance of the correction procedure to the semi-variogram parameters, true validation was performed at various combinations of the types of the model, n_g, n_r, L_g, and L_r. The resulting performance measures were then examined in the form of contour plots for all pairwise combinations of the parameters. The analysis indicate that: 1) the semi-variogram models of gauge and radar rainfall should be of the same type (e.g. exponential, gaussian, or spherical) even if it may not best-fit both experimental semivariograms, and 2) the performance of the proposed procedure is much more sensitive to the correlation scale parameters, L_g and L_t, than to the normalized nugget effect parameters, n_g and n_r. The latter point is illustrated in Fig 4, in which the contour plots of RMSE are shown as functions of L_g and n_g for cool and warm seasons.

In light of the above observations, we chose the exponential semi-variogram model for both radar and gauge rainfall, set $n_g = n_r \approx 0$, and made true-validation runs at various combinations of L_g and L_r , each parameter assuming values of 1, 3, 5, 7, and 9 (HRAP). Figs 5 and 6 show the resulting contour plots of RATIO and RMSE of the hourly local bias-corrected estimates for all amounts of gauge rainfall and for amounts exceeding 25.4 mm as functions of L_g and L_r for cool and warm seasons, respectively. All figures are based on $N_t = 8$, which was found to provide an appropriate level of temporal smoothing for estimation of daily amounts in both seasons. Figs 5a

and 5b indicate that, in the cool season, the minimum RMSE-producing combinations of L_g and L_r are not necessarily the best from the unbiasedness point of view. Because the overriding objective of bias correction is to achieve RATIO \approx 1, the optimal combination of L_g and L_r should be located where RMSE is the smallest in Figs 5b and 5d given RATIO \approx 1 in Fig 5a. This results in the cool-season setting of $(L_g,L_r)\approx(5,3)$ (HRAP). For the warm season, the minimum RMSE-producing combinations of L_g and L_r in Figs 6b and 6d also produce RATIO \approx 1 in Fig 6a. Hence, the optimal parameter combination may be found where RMSE is the smallest in Figs 6b and 6d. This results in the warm-season setting of $(L_g,L_r)\approx(1,1)$ (HRAP), though $(L_g,L_r)\approx(3,3)$ (HRAP) is also a possibility.

Note in Figs 5c and 6c that, even at these 'optimal' parameter settings, significant biases still exist for estimation of large rainfall amounts in both seasons. These conditional biases arise from lack of sampling and from the use of a linear estimator even though rainfall has a highly skewed distribution. The magnitude of the conditional bias at ABRFC should in reality be much smaller than that suggested by Figs 5c and 6c because the full gauge network, rather than the ½ network used in the true validation above, is available operationally for estimation of local bias. Conceptually, use of nonlinear estimation should reduce the conditional bias (see, e.g., Seo 1996). Given, however, that nonlinear estimation requires parameterization, be it implicit (as, e.g., in Artificial Neural Networks) or explicit (as, e.g., in disjunctive (Journel and Huijbregts 1978) or indicator kriging (Deutsch and Journel 1992)), of statistical moments that are of higher order than that of the bias itself, it is seen unlikely that nonlinear estimation can be used successfully for real-time bias correction.

With the (near-) optimal parameter settings identified above, true validation was performed using the two-year validation data set. Fig 7 shows the scatter plots of daily amounts of the raw, mean field bias-corrected, and local bias-corrected radar rainfall estimates against gauge rainfall for the cool and warm seasons. They indicate that the local bias-corrected estimates are generally better, particularly in the cool season, than the mean field bias-corrected estimates. A number of gross overestimates seen in Fig 7b are associated with a bright band event for which mean field bias was estimated mostly from the radar-gauge pairs in the far range, where radar tends to severely underestimate rainfall due to sampling of ice particles above the melting layer. Application of the mean field bias everywhere in the effective coverage of the radar consequently results in gross overestimation of mean field bias-corrected radar rainfall in the mid-range, where radar tends to overestimate rainfall due to bright band enhancement.

Table 2 shows the summary statistics from the validation run. They are tabulated according to the season (cool and warm), to the temporal scale of aggregation (hourly and daily), and to the amount of rainfall being estimated (all amounts and large amounts only). The statistics confirm the observations made in Fig 7 that the local bias-corrected estimates are generally superior, particularly in the cool season, to the mean field bias-corrected estimates. The rather poor performance of the mean field bias correction procedure in removing the mean bias in the cool season is attributed to a number of prominent bright band events in the validation period. The local bias-corrected estimates, on the other hand, are free of mean bias in both cool and warm seasons. Note that local bias correction also significantly improves estimation skill over mean field bias correction in the cool season. For example, the percent reductions in MSE of daily amounts by the local bias- and mean field bias-corrected estimates over the raw radar

rainfall are about 46 and 27 (%), respectively, in the cool season, and 31 and 26 (%), respectively, in the warm season.

In many parts of the country, the operational (near-) real-time hourly rain gauge network is sparser than the ½ network on which the results presented above are based. To evaluate the performance of the correction procedure under a sparser network, validation runs were also made using the 1/4 network (see Section 7). The semi-variogram parameters used were the same as those used in the ½ network case. The results are summarized in Table 3, which is to be compared with Table 2. They indicate that, under a sparser network, the performance of the local bias correction procedure generally deteriorates, but remains superior, particularly in the cool season, to mean field bias correction.

4.9 Conclusions and Future Research Recommendations

A procedure for real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements is described. Based on operational experience with the existing bias correction procedures used in the National Weather Service (NWS), the proposed procedure is intended to be a generalized local bias estimator that may be used under varying conditions of gauge network density and types of rainfall. For parameter estimation and evaluation of the procedure, WSR-88D Digital Precipitation Array (DPA) products and operational hourly rain gauge data were used from the Arkansas-Red Basin River Forecast Center (ABRFC) area covering May 1, 1996, through April 30, 1999.

Parameter estimation was carried out via true validation using data from the first year in a series of sensitivity analyses. The results indicate that 1) the proposed procedure is most sensitive to the spatial correlation scales of gauge and radar rainfall, and 2) as such, site-specific sensitivity analysis or parameter optimization is necessary to pin down the optimal parameter settings and to ascertain the nature of the dependence of the conditional (i.e. on the amount of rainfall being estimated) performance of the procedure on the parameter settings.

Evaluation of the proposed procedure was carried out via true validation using data from the last two years. The results indicate that the proposed procedure is generally superior, particularly in the cool season, to the mean field bias correction procedure (Seo et al. 1999): the proposed procedure is very effective in removing mean bias (i.e. bias averaged over all amounts of rainfall) for both cool and warm seasons, and reduces the mean square error (MSE) of daily rainfall, relative to that of the raw radar rainfall estimates, by at least 34 to 46 and 23 to 31 (%) (16 to 27 and 17 to 26 (%) for mean field bias correction) in the cool and warm seasons, respectively, over the range of rain gauge network density examined in this work.

The sensitivity of the proposed procedure to the parameters suggests that the biggest improvement is most likely to come from further stratification of the correlation scale parameters (including accounting of anisotropy) beyond that based on seasonality alone. Identification of the attributes that lead to more skillful stratification of the parameters and determination of the space-time scale at which such stratification may remain practical warrant further investigation.

APPENDIX A Local Bias Correction Procedure Used in Process 1 (P1)

In the simplest terms, the local bias correction procedure used in P1 may be described by the following:

where

$$\beta_{i} = \begin{cases} 1 & \text{if } g_{i}/r_{i} > \beta_{t} \\ g_{i}/r_{i} & \text{if } g_{i}/r_{i} \leq \beta_{t} \end{cases}$$
(A2)

$$\delta_{i} = \begin{cases} g_{i} - r_{i} & \text{if } g_{i} / r_{i} > \beta_{t} \\ 0 & \text{if } g_{i} / r_{i} \leq \beta_{t} \end{cases}$$
(A3)

In Eq.(A1), r_c denotes the bias-corrected radar rainfall (mm), r_0 denotes the raw radar rainfall at the bin centered at u_0 (mm), w_i denotes the weight given to the radar-gauge pair at the i-th vertex in the triangle of radar-gauge pairs that encloses u_0 , β_i denotes the multiplicative sample bias from the i-th radar-gauge pair, and δ_i denotes the additive sample bias from the i-th radar-gauge pair. In Eqs.(A2) and (A3), g_i and r_i denote the gauge rainfall measurement (mm) and the collocating radar rainfall estimate (mm), respectively, at the i-th vertex in the enclosing triangle, and β_i , an adaptable parameter, denotes the threshold for the multiplicative sample bias, g_i/r_i , i=1,2,3. The neighboring radar-gauge pairs are identified by "triangulation," which connects all available radar-gauge pairs into a mesh of triangles. The weights, w_i , i=1,2,3, sum to unity and are proportional to the distance to the neighboring radar-gauge pairs in the enclosing triangle.

APPENDIX B

Identity Between the Fisher Estimate of Eq.(10) and the Ordinary Kriging Estimate

Here we show that, under Eq.(13), the Fisher estimate (Schweppe 1973) of R_0 in Eq.(10) (the time index has been dropped for notational brevity), i.e., the unbiased minimum-error-variance estimate of R_0 given that nothing is known a priori about its statistics, is identical to the ordinary kriging estimate. The Fisher estimate of R_0 , R_0^* , and the associated estimation variance, Σ_0 , are given by (Schweppe 1973):

$$\Sigma_0 = (H^T \Psi^{-1} H)^{-1}$$
(B1)

$$R_0^* = \Sigma_0 H^T \Psi^{-1} Z \tag{B2}$$

The terms in the above expressions are defined in Section 4. We rewrite Eq.(B2) as;

$$R_0^*$$
= $(\Lambda H)^{-1}\Lambda Z$
= WZ
(B3)

where

$$\Lambda = \mathbf{H}^{\mathsf{T}} \mathbf{\Psi}^{-1} \tag{B5}$$

$$\begin{array}{c}
n \\
W = \Lambda/\Sigma \lambda_i \\
i=1
\end{array}$$
(B6)

In the above, n denotes the number of neighbors used in the estimation and λ_i denotes the i-th entry in Λ . Post-multiplying Ψ to Eq.(B5) and taking transpose, we have:

$$\Psi^{\mathsf{T}}\Lambda^{\mathsf{T}} = \mathsf{H} \tag{B7}$$

With Eq.(13), (B7) may be written as:

where ψ_{ij} denotes the ij-th entry in the error covariance matrix Ψ . The above linear system may be rearranged into the following ordinary kriging system:

$$\begin{bmatrix} \psi_{11} \dots \psi_{1n} \\ \vdots \\ \vdots \\ \psi_{n1} \dots \psi_{nn} \end{bmatrix} \begin{bmatrix} W_1 \\ \vdots \\ W_n \end{bmatrix} - \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \mu = \begin{bmatrix} \psi_{10} \\ \vdots \\ \vdots \\ \psi_{n0} \end{bmatrix}$$
(B9)

In the above, w_i denotes the i-th entry in the weight vector W in (B6) (note that by definition, w_i , i=1,...,n, sum to unity), and the Lagrange multiplier μ is given by:

APPENDIX C

Non-Recursive Form of the Exponential Smoothing Solution

Given the initial conditions $r_{00|0}=0$ and $\Sigma_{r00|0}^{-1}=0$, Eqs.(11) and (12) may be rewritten in the following non-recursive forms:

$$\begin{array}{c|c}
 k \\
 r_{0k|k} = \sum_{i=1}^{k} \left[\omega_{i} \pi \left\{ 1 - \omega_{j} \right\} \right] R_{0i}^{*} \\
 i = 1 \end{array} (C1)$$

$$\sum_{0k|k}^{-1} = \sum_{i=1}^{k} e^{-(k-i)/\alpha} \sum_{0i}^{-1}$$
(C2)

In the above, Σ_{0i} and R_{0i}^* are given by Eqs.(B1) and (B2), respectively, and the weight given to R_{0i}^* , ω_i , is given by:

$$\omega_{i} = \Sigma_{0i}^{-1} / \Sigma_{0i \mid i}^{-1} \tag{C3a}$$

$$= \sum_{0i}^{-1} / (e^{-1/\alpha} \sum_{0(i-1)|(i-1)}^{-1} + \sum_{0i}^{-1})$$
 (C3b)

Eqs.(C1) through (C3) state the exponential smoothing solution for R_{0k} , $r_{0k|k}$, is a linearly weighted average of R_{0i}^* , i=1,...,k, i.e., a temporal average of ordinary kriging estimates at all hours up to and including the current hour as linearly-weighted according to the 'age' of and the magnitude of the error variance associated with the estimates. Note that, because the weights sum to unity, i.e., $\Sigma_{i=1}^k \omega_i \pi_{i=i+1}^k \{1-\omega_i\}=1$, $r_{0k|k}$ is also unbiased in the mean sense.

APPENDIX D

Exponential Smoothing of Bin-Averaged Gauge Rainfall Using (Point) Gauge Measurements

The observation equation for the bin-average gauge rainfall centered at u_0 at hour k, G_{0k} , is given by:

$$Z_{gk} = H_k G_{0k} + V_{gk}$$
 (D1)

where the observation vector, Z_{gk} , the structure vector, H_k , the unknown bin-averaged gauge rainfall centered at u_0 , G_{0k} , and the error vector, V_{gk} , are given by $[g_{1k},g_{2k},...,g_{(nk)k}]^T$, $[1,1,...,1]^T$, $\|A_0\|^{-1}\int_{A0}P_k(u)du$, and $[G_{1k}-\|A_0\|^{-1}\int_{A0}P_k(u)du$, $G_{2k}-\|A_0\|^{-1}\int_{A0}P_k(u)du$,..., $G_{(nk)k}-\|A_0\|^{-1}\int_{A0}P_k(u)du$, respectively. In the above, g_k denotes the i-th rain gauge measurement at hour k, and P(u) denotes the point hourly gauge rainfall at location u within the bin. The ij-th entry in the error covariance matrix of V_{gk} , Ψ_{gk} , is then given by:

 ψ_{gijk}

$$= Cov[G_{ik} - ||A_0||^{-1} \int_{A_0} P_k(u) du, G_{ik} - ||A_0||^{-1} \int_{A_0} P_k(u) du]$$
(D2a)

$$= \|A_0\|^{-2} \int_{A_0} \int_{A_0} Cov[P_k(u), P_k(v)] dudv$$

$$- \|A_0\|^{-1} \int_{A_0} Cov[G_{ik}, P_k(u)] du$$

$$-\|A_0\|^{-1}\int_{A_0} Cov[P_k(u),G_{ik}]du$$

$$+ \operatorname{Cov}[G_{ik}, G_{ik}]$$
 (D2b)

$$= \big\|A_0\big\|^{\text{-}1}\big\}_{\text{A0}} \, \gamma[G_{ik},\!P_k(u)]du + \big\|A_0\big\|^{\text{-}1}\big\}_{\text{A0}} \, \gamma[P_k(u),\!G_{jk}]du$$

$$- \|A_0\|^{-2} \int_{A_0} \gamma[P_k(u), P_k(v)] du dv - \gamma[G_{ik}, G_{ik}]$$
(D2c)

where Cov[] and $\gamma[]$ denote the covariance and the semi-variogram, respectively.

APPENDIX E

Effect of Using (Point) Gauge Measurements, in Lieu of Areally-Averaged Gauge Rainfall, on Mean Square Error Calculations

The mean square error of raw radar rainfall estimates with respect to the (point) gauge measurements, $MSE[R_r-P(u)]$ is an estimate of $E[\{R_r-P(u)\}^2]$, where R_r denotes the raw radar rainfall and P(u) denotes the gauge rainfall at some point u within the radar bin. We may rewrite $E[\{R_r-P(u)\}^2]$ as:

$$E[\{R_r-P(u)\}^2]=E[\{R_r-G+G-P(u)\}^2]$$
(E1)

In the above, G denotes the areally averaged gauge rainfall over the radar bin, i.e., $G = \|A_0\|^{-1} \int_{A_0} P(u) du$, where A_0 denotes the area of the bin. Then, under the assumption that the estimation error in the raw radar rainfall estimate, R_r -G, and the sampling error in the (point) gauge measurement due to micro-scale variability of rainfall, G-P(u), are linearly independent (Drake 1967: see Ciach and Krajewski 1999 for justification), we may write:

$$E[{R_r-P(u)}^2] \approx E[(R_r-G)^2] + Var[G-P(u)]$$
 (E2)

Likewise, under the similar assumption that R_c -G and G-P(u) are linearly independent, where R_c denotes the bias-corrected radar rainfall, we may write:

$$E[\{R_c-P(u)\}^2] \approx E[(R_c-G)^2] + Var[G-P(u)]$$
(E3)

From Eqs.(E2) and (E3), we then have:

$$E[\{R_r-P(u)\}^2]-E[\{R_c-P(u)\}^2] \approx E[(R_r-G)^2]-E[(R_c-G)^2]$$
 (E4)

$$\begin{split} & E[\{R_r - P(u)\}^2] - E[\{R_c - P(u)\}^2] & E[(R_r - G)^2] - E[(R_c - G)^2] \\ & - E[\{R_r - P(u)\}^2] & E[(R_r - G)^2] \end{split}$$
 (E5)

In other words, the true reduction in mean square error based on bin-averaged gauge rainfall, $MSE[R_r-G]-MSE[R_c-G]$, is approximately the same as the apparent reduction, $MSE[R_r-P(u)]-MSE[R_c-P(u)]$, based on (point) rain gauge measurements (Eq.(E4)), and the true percent reduction in mean square error based on bin-averaged gauge rainfall is larger than the apparent percent reduction based on (point) rain gauge measurements (Eq.(E5)).

REFERENCES

- Ahnert, P., W. Krajewski, and E. Johnson, 1986: Kalman filter estimation of radar-rainfall field bias. Preprints, 23rd Conf. On Radar Meteor., Snowmass, CO, Amer. Meteor. Soc., JP33-JP37.
- Anagnostou, E. N., W. F. Krajewski, D.-J. Seo, and E. R. Johnson. 1998: Mean-field rainfall bias studies for WSR-88D. J. Hydrol. Eng., 3, 149-159.
- Andrieu, H., and J. D. Creutin, 1995: Identification of vertical profile of radar reflectivity for hydrological applications using an inverse method. Part I: Formulation. J. Appl. Meteor., 34, 225-239.
- Bastin, G, B. Lorent, C. Duque, and M. Gevers, 1984: Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations. Water Resour. Res., 20(4), 463-470.
- Bras, R. L., and I. Rodriguez-Iturbe, 1985: Random functions and hydrology, Addison-Wesley, MA, 559pp.
- Breidenbach, J. P., D.-J. Seo, P. Tilles, and K. Roy, 1999: Accounting for radar beam blockage patterns in radar-derived precipitation mosaics for River Forecast Centers. Preprint Volume, 15th Conf. On IIPS, Amer. Meteor. Soc., Dallas, TX.
- Breidenbach, J. P., D.-J. Seo, P. Tilles, and C. Pham, 2000: Seasonal variation in multiradar coverage for WSR-88D precipitation estimation in a mountainous region. Preprint Volume, Conf. On Precipitation Extremes: Prediction, Impacts, and Responses, Amer. Meteor. Soc., Albuquerque, NM.
- Ciach, G. J., and W. F. Krajewski, 1999: On the estimation of radar rainfall error variance. Adv. Water Resour., 22, 585-595.
- Collier, C. G., 1986: Accuracy of rainfall estimates by radar, Part I: Calibration by telemetering raingauges. J. Hydrol. 83, 207-223.
- Deutsch, C. V., and Journel, A. G., 1992. GSLIB Geostatistical software library and user's guide. Oxford University Press, 340pp.
- Doviak, R. J., and D. S. Zrnic, 1984: Doppler radar and weather observations. Academic Press, 458pp.
- Drake, A. W., 1967: Fundamentals of applied probability theory. McGraw-Hill, Inc., pp283.

- Fabry, F., A. Bellon, M. R. Duncan, and G. L. Austin, 1994: High resolution rainfall measurements by radar for very small basins: the sampling problem reexamined. J. Hydrol. 161, 415-428.
- Fulton, R. A., J. P. Breidenbach, D.-J. Seo, and D. A. Miller, 1998: The WSR-88D rainfall algorithm. Weather and Forecasting, 13, 377-395.
- Hudlow, M. D., 1988: Technological developments in real-time operational hydrologic forecasting in the United States. J. Hydrol., 102, 69-92.
- Journel, A. G., and Ch. J. Huijbregts, 1978: Mining geostatistics, Academic Press, 600pp.
- Klazura, G., and D. Imy, 1993: A description of the initial set of analysis products available from the NEXRAD WSR-88D system. Bull. Amer. Meteor. Soc., 74, 1293-1311.
- Lebel, T., J. P. Larborde, 1988: A geostatistical approach to areal rainfall statistical assessment. Stoch. Hydrol. Hydraul., 2(4), 245-261.
- Lebel, T., G. Bastin, C. Obled, and J. D. Creutin, 1987: On the accuracy of areal rainfall estimation: A case study. Water Resour. Res., 22(11) 2123-2134.
- O'Bannon, T., 1997: Using a 'terrain-based' hybrid scan to improve WSR-88D precipitation estimates. Preprints, 28th Conf. On Radar Meteorology, Austin, TX, Amer. Meteor. Soc., 506-507.
- Rodriguez-Iturbe, I. and Mejia, 1974: The design of rainfall network in time and space. Water Resour. Res., 10(4), 713-728.
- Rouhani, S. and D. E. Myers, 1990: Problems in space-time kriging of geohydrological data. Math. Geol., 22(5), 611-623.
- Rouhani, S. and H. Wackernagel, 1990: Multivariate geostatistical approach to space-time data analysis. Water Resour. Res., 26(4), 585-591.
- Saffle, R. E. and L. D. Johnson, 1999: NEXRAD Open Systems Progress and plans. Preprint Volume, 15th Conf. On IIPS, Amer. Meteorol. Soc., Dallas, TX.
- Schweppe, F. C., 1973: Uncertain Dynamic Systems. Prentice-Hall, Englewood Cliffs, NJ, 563pp.
- Seo, D.-J., 1996: Nonlinear estimation of spatial distribution of rainfall An indicator cokriging approach. Stoch. Hydrol. Hydraul., 10, 127-150.
- Seo, D.-J., 1998: Real-time estimation of rainfall fields using rain gauge data under fractional coverage conditions. J. Hydrol., 208, 25-36.
- Seo, D.-J., and Smith, J. A., 1994: On space-time estimation of rainfall fields using rain gauge data. Technical Note, NWS/OH/HRL, Silver Spring, MD (available upon request).
- Seo, D.-J., R. A. Fulton, and J. P. Breidenbach, 1997: Final Report, Interagency MOU among the NEXRAD Program, WSR-88D OSF, and NWS/OH/HRL, NWS/OH/HRL, Silver Spring, MD (available upon request).
- Seo, D.-J., J. P. Breidenbach, and E. R. Johnson, 1999: Real-time estimation of mean field bias in radar rainfall data, J. Hydrol., 223, 131-147.
- Seo, D.-J., J. P. Breidenbach, R. A. Fulton, D. A. Miller, and T. O'Bannon, 2000: Real-time adjustment of range-dependent bias in WSR-88D rainfall data due to nonuniform vertical profile of reflectivity. J. Hydrometeorol., 1(3), 222-240.
- Smith, J. A., and W. F. Krajewski, 1991: Estimation of the mean field bias of radar rainfall estimates. J. Appl. Meteor., 30, 397-412.

- Steiner, M., J. A. Smith, S. J. Burgess, C. V. Alonso, and R. W. Darden, 1999: Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. Water Resour. Res., 35(8), 2487-2503.
- Tzafestas, S. G., 1978: Distributed parameter systems. Ch. 3, Distributed parameter state estimation. W. H. Ray and D. G. Lainiotis, Eds., Dekker, New York.
- Wilson, J., and E. Brandes, 1979: Radar measurement of rainfall A summary. Bull. Amer. Meteor. Soc., 60, 1048-1058.
- Young, C. B., A. A. Bradley, W. F. Krajewski, and A. Kruger, 2000: Evaluation NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. J. Hydrometeorol., 1(3), 241-254.

Table 1. Climatological semi-variogram parameters.

season	sensor	type	L ¹ (HRAP)	C^2 (mm ²)	$\mu^3 (mm^2)$
cool	radar gauge	sp ⁴ ex ⁶	12.10 ⁵ 5.02	26.9 38.8	0.0 0.0
warm	radar gauge	ex ex	3.36 3.50	60.2 76.4	0.0 0.0

¹ spatial correlation scale

Table 2a. Performance measures for hourly estimates (gauge rainfall > 0, $N_t = 8$, $\frac{1}{2}$ network).

season	type	RATIO	RMSE ¹	CORR	MAXEU ²	MAXEO ²
Cool	raw³	1.35	3.49	0.71	36.3	32.4
	mfb⁴	0.93	3.74	0.75	33.9	51.4
	lb⁵	1.00	3.33	0.78	42.3	44.0
Warm	raw	1.07	4.48	0.73	48.0	47.1
	mfb	1.03	4.22	0.76	39.8	50.6
	lb	1.00	4.22	0.77	33.4	56.2

¹ in mm²

² sill

³ nugget effect

⁴ spherical model

⁵ corresponds to 6.13 (HRAP) in the exponential model

⁶ exponential model

² in mm

³ uncorrected radar rainfall estimate

⁴ mean field bias-corrected radar rainfall estimate

⁵ local bias-corrected radar rainfall estimate

Table 2b. Performance measures for hourly estimates (gauge rainfall > 25.4 mm, N_t =8, $\frac{1}{2}$ network).

season	type	RATIO	RMSE	CORR	MAXEU	MAXEO
Cool	raw	1.62	16.0	0.51	36.3	13.0
	mfb	1.13	12.9	0.44	33.9	43.9
	lb	1.12	12.2	0.51	42.3	36.5
Warm	raw	1.43	15.6	0.49	48.0	40.2
	mfb	1.43	14.7	0.49	39.8	32.3
	lb	1.35	14.5	0.53	33.4	56.2

Table 2c. Performance measures for daily estimates (gauge rainfall > 0, $N_t=8$, $\frac{1}{2}$ network).

season	type	RATIO	RMSE	CORR	MAXEU	MAXEO
Cool	raw mfb lb	1.35 0.93 1.00	9.33 7.98 6.86	0.82 0.87 0.90	107.1 82.9 74.0	64.7 166.0 65.3
Warm	raw mfb	1.08	8.25 7.08	0.80 0.86	70.5 50.7	109.6 50.6
	lb	1.00	6.87	0.87	46.2	55.4

Table 2d. Performance measures for daily estimates (gauge rainfall > 50.8 mm, N_t =8, $\frac{1}{2}$ network).

season	type	RATIO	RMSE	CORR	MAXEU	MAXEO
Cool	raw	1.70	40.5	0.68	107.1	64.7
	mfb	1.09	22.6	0.76	82.9	71.4
	lb	1.12	22.9	0.77	74.9	47.4
Warm	raw	1.33	30.2	0.51	70.5	109.6
	mfb	1.23	21.4	0.64	50.7	26.0
	lb	1.16	20.4	0.64	46.2	55.4

Table 3a. Performance measures for hourly estimates (gauge rainfall > 0, $N_t=16$, 1/4 network).

season	type	RATIO	RMSE	CORR	MAXEU	MAXEO
Cool	raw	1.34	3.48	0.70	38.5	36.8
	mfb	0.89	3.98	0.72	32.1	49.5
	lb	0.98	3.56	0.74	43.9	46.0
Warm	raw	1.08	4.52	0.73	41.9	47.2
	mfb	1.05	4.41	0.74	39.1	59.5
	lb	1.02	4.36	0.75	38.7	62.7

Table 3b. Performance measures for hourly estimates (gauge rainfall > 25.4 mm, N_t =16, 1/4 network).

season	type	RATIO	RMSE	CORR	MAXEU	MAXEO
Cool	raw	1.59	15.8	0.47	38.5	13.8
	mfb	1.10	13.0	0.44	32.1	40.1
	lb	1.19	12.7	0.47	43.9	35.7
Warm	raw	1.44	15.5	0.46	41.9	40.2
	mfb	1.46	15.5	0.43	39.1	37.8
	lb	1.43	14.7	0.48	38.7	39.3

Table 3c. Performance measures for daily estimates (gauge rainfall > 0, $N_t=16$, 1/4 network).

season	type	RATIO	RMSE	CORR	MAXEU	MAXEO
	raw	1.34	9.14	0.82	107.1	64.7
Cool	mfb	0.89	8.37	0.86	88.8	108.6
	lb	0.98	7.43	0.88	86.5	70.6
	raw	1.08	8.26	0.80	117.4	109.6
Warm	mfb	1.06	7.51	0.84	76.7	59.5
	lb	1.02	7.25	0.85	75.9	62.7

Table 3d. Performance measures for daily estimates (gauge rainfall > 50.8 mm, N_t =16, 1/4 network).

season	type	RATIO	RMSE	CORR	MAXEU	MAXEO
Cool	raw	1.68	38.7	0.67	107.1	64.7
	mfb	1.06	23.1	0.74	88.8	68.2
	lb	1.14	22.3	0.78	86.5	65.3
Warm	raw	1.37	30.5	0.47	117.4	109.6
	mfb	1.30	23.8	0.58	76.7	41.3
	lb	1.27	22.3	0.61	75.9	49.4

List of Figure Captions

- Fig 1 Radar (call letters) and rain gauge locations (dots) within and in the vicinity of the ABRFC's service area (dotted line)
- Fig 2 a) experimental semi-variograms (see text for explanation of symbols) of cool-season radar rainfall and the best weighted least-squares fit (solid line), b) same as a), but for cool-season gauge rainfall, c) same as a), but for warm-season radar rainfall, and d) same as a), but for warm-season gauge rainfall
- Fig 3 a) RATIO and RMSE of the hourly raw ('r'), mean field bias-corrected ('m'), and local bias-corrected ('1' through '6': see text for explanation) radar rainfall estimates in the cool season, b) same as a), but in the warm season, c) cool- and warm-season CORRE of the hourly raw ('r'), mean field bias-corrected ('m'), and local bias-corrected ('1' through '6': see text for explanation) radar rainfall estimates, d) same as a), but for MAXEU and MAXEO, e) same as b), but for MAXEU and MAXEO
- Fig 4 a) example of the contour plot of RMSE of hourly local bias-corrected radar rainfall estimates as a function of $L_{\rm g}$ and $n_{\rm g}$ for the cool season, b) same as a), but for the warm season
- Fig 5 a) contour plot of RATIO of hourly local bias-corrected radar rainfall estimates for all amounts of gauge rainfall as a function of $L_{\rm g}$ and $L_{\rm r}$ for the cool season, b) same as a), but for RMSE, c) same as a), but for gauge rainfall amounts exceeding 25.4 (mm), d) same as b), but for gauge rainfall amounts exceeding 25.4 (mm)
- Fig 6 a) same as Fig 5a, but for the warm season, b) same as Fig 5b, but for the warm season, c) same as Fig 5c, but for the warm season, d) same as Fig 5d, but for the warm season
- Fig 7 a) scatter plot of daily rainfall amounts in the cool season between gauge and raw radar rainfall, b) same as a), but for the warm season, c) same as a), but between gauge and mean field bias-corrected radar rainfall, d) same as c), but for warm season, e) same as a), but between gauge and local bias-corrected radar rainfall, f) same as e), but in the warm season





































